

An Interactive Lifelog Search Engine for LSC2018

Adrià Alsina & Xavier Giró
Universitat Politècnica de Catalunya
Barcelona, Spain

adria.alsinatorra2@mail.dcu.ie, xavier.giro@upc.edu

Cathal Gurrin
Dublin City University
Dublin, Ireland
cathal.gurrin@dcu.ie

ABSTRACT

In this work, we describe an interactive lifelog search engine developed for the LSC 2018 search challenge at ACM ICMR 2018. The paper introduces the four-step process required to support lifelog search engines and describes the source data for the search engine as well as the approach to ranking chosen for the iterative search engine. Finally the interface used is introduced before we highlight the limits of the current prototype and suggest opportunities for future work.

CCS CONCEPTS

• Information systems → Digital libraries and archives; Search interfaces;

KEYWORDS

Lifelogging, Interactive Search Engine, Information Retrieval

ACM Reference Format:

Adrià Alsina & Xavier Giró and Cathal Gurrin. 2018. An Interactive Lifelog Search Engine for LSC2018. In *LSC'18: The Lifelog Search Challenge, June 11, 2018, Yokohama, Japan*. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3210539.3210546>

1 INTRODUCTION

Lifelogging is a new concept and has only recently attracted the attention of the research community [7]. One standard definition of lifelogging states that it is *"the action or practice of making a continuous record of one's daily activities by means of a digital device or computer application"*. The growing research interest in lifelogging has been facilitated by the market-availability of a range of lifelogging and quantified self devices that can digitally capture a wide range of life activities, from wearable cameras for all-of-lifelogging to the more targeted health and wellness devices commonly referred to as quantified self devices [9].

With such a range of lifelogging devices, one can easily gather, in an automated manner, a wide range of data about the life experience of the individual, such as image data from smartphones or wearable cameras, audio recordings, location data, biometric data, to name only a few.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

LSC'18, June 11, 2018, Yokohama, Japan

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5796-8/18/06...\$15.00

<https://doi.org/10.1145/3210539.3210546>

The increase of interest in lifelogging has resulted in many different research challenges being developed, such as the NTCIR lifelog task Semantic Access Task (LSAT) of the the NTCIR-12 challenge [6] and the ImageCLEF [8] lifelog task. This LSC lifelog task posed unique challenges due to the interactive nature of the task, hence we felt that our retrieval approaches would be suitable for this task. The interactive nature of the challenge means that any system developed for the LSC would need to be optimised to select or locate relevant content from a comparatively small archive of lifelog data, which in this case, consisted of multimedia data from wearable cameras, biometric data from smartwatches, human activity data from smartphones and computer usage data to identify information access activities.

Without there having been similar tasks in the past, there are not many interactive retrieval engines that were developed for lifelog data previously. The seminal MyLifeBits [5] project at Microsoft also developed a lifelog retrieval engine based on an underlying database system, which is generally regarded as the first lifelog retrieval system. The LEMoRe [2], an interactive lifelog retrieval engine, developed in the context of the Lifelog Semantic Access Task (LSAT) of the the NTCIR-12 challenge [6]. LEMoRe integrated classical image descriptors with high-level semantic concepts extracted by Convolutional Neural Networks [13], powered by a graphical user interface that uses natural language processing to process a user's query. Finally, Doherty et al. [4], developed an interactive event-driven lifelog browser for visual lifelog data that segmented days into events, based on analysis of visual and sensor data, organising and linking events together in a single diary-style interface. Extending this work, an interactive faceted lifelog search engine [3] was developed that allowed the user to select a number of contextual factors in order to find an event of interest from an archive that spanned a number of years.

Building on experiences from this past work, we have developed the experimental search engine which forms the main contribution of this paper. The general approach taken for this prototype is to develop an indexing and filtering tool, utilising an SQL database [5], with a customised interface that is designed to support fast free-text and faceted search (faceted search based on [3]). The difference between this search engine and the preceding efforts is that free text search, different facets, and the interface presentation that considers the temporal context of the lifelogger when presenting ranked or filtered results.

2 A FOUR-STEP APPROACH TO INTERACTIVE LIFELOG RETRIEVAL

In designing our lifelog retrieval engine, we separate the task into four challenges:

- **Data Selection and Enrichment.** Lifelogging typically involves multiple sensors, as did the LSC dataset. The first task is to choose, process and align these data files so that they are temporally arranged.
- **Multi-faceted query interface.** The multi-modal nature of the LSC dataset naturally lends itself to multi-faceted query generation, so we have develop a query mechanism that allows a user to enter queries based on free-text or facets, such as date/time and place.
- **Ranking Engine.** For retrieving potentially relevant content for the user using an appropriate ranking technique.
- **Presentation of a Result List.** Given the temporal nature of lifelog data, it is sensible to consider temporal organisation of the result list, since the ranking employed will hugely influence how effective the interactive retrieval system can be.

We will now explain how we solved each of these four challenges.

2.1 Data Selection and Enrichment

The LSC lifelog dataset consists of a wide array of multimodal data of an individual's life experience captured via wearable and software sensors. The dataset was gathered over a period of 27 days by a single individual and it includes:

- **Multimedia data:** Represents about 1,500 images per day taken by a wearable camera. So, if we have 27 days of data it means that we have a database of images about 40,500 images. These images are accompanied by the output of a concept detector which identifies the dominant visual concepts in every image. In addition, the music listening history of the lifelogger is included, though we did not make use of it in this prototype.
- **Biometric data:** It is formed for biometric data like heart rate, galvanic skin response, calorie burn and steps stored almost every minute of the day. Also we have data about blood pressure and blood sugar daily. In this search engine, we utilised the heart rate data and segmented it into three categories (resting, normal, and physically active).
- **Human activity data:** Consists in information about the semantic locations visited, the physical activities, the daily mood and a diet log made of manual logging of photos. We focused on the semantic locations and physical activities in the search engine.
- **Computer Usage:** Represents a vector of every minute stored data filtered using blacklist, anonymised and then stemmed. We did not use this data in this work.

In addition, each data item described above is timestamped and this was used to time-align the data for indexing. We also utilised this time information to create a semantic time annotation for every moment in the collection, such as day-of-the-week or time-of-the-day (morning, afternoon, evening).

The image formed the main unit of retrieval for this system and every image was annotated with the appropriate enriched semantic content from the dataset. These images were later selected for display to the user during the interactive retrieval process.

2.2 Multi-faceted query interface

Given the interactive nature of the LSC exercise, the design of the interface, both for rapid querying and efficient browsing of result sets becomes the most important task. Borrowing from the standard WWW-interface for faceted search systems (e.g. hotel booking or flight booking), we designed the interface with two sections, as in *Figure 1*. On the left side the query panel is displayed which contains the faceted and free-text query elements. On the right side is the result display panel.

Given the fact that the number of unique annotated terms in the collection could be very large (especially in terms of visual concepts and locations), it was decided to replace the typical drop-down list with a free-text search box on the query panel. This search box matched against content in any of the underlying data sources and is hoped to simplify the use of the system for the user. Given a limited lexicon of available search terms, the search box offered auto-completion of input terms, which assists the user in formulating free text queries. In addition to the free-text search, some more conventional faceted search filter options were available:

- **Day-of-the-Week selector** to filter content by the day the event represented in the topic occurred.
- **Calendar selector** to choose an exact date.
- **Moment-of-the-Day selector** to filter content that occurs at certain times of the day.
- **Place selector** to allow filtering into indoor/outdoor events.
- **Heart-rate selector** to choose the various degrees of physical activity of the lifelogger.

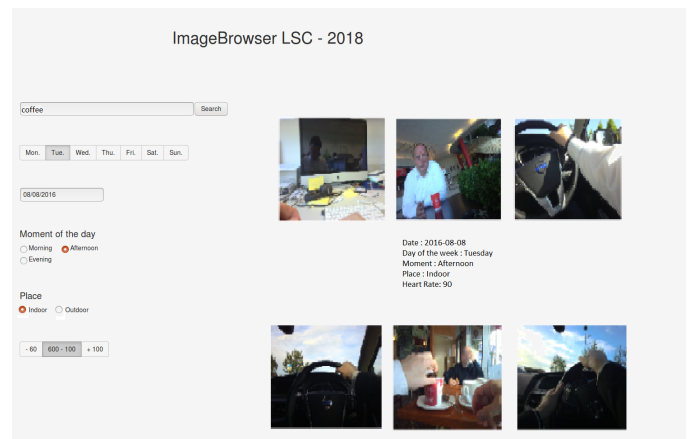


Figure 1: Display of the images.

In terms of technologies the client interface is developed using the JavaScript library React.js. The server-side data storage is handled by SQL Server and communication is via the python library Flask. The technologies are chosen so as to optimise the speed of handing every faceted query change and the system has been designed to be responsive to any facet updates.

2.3 Ranking Engine

The ranking engine used in this interactive lifelog search tool integrates TF-IDF ranking for the free-text search as well as the faceted

filtering mechanism. TF-IDF is a statistical approach to information retrieval that is intended to reflect how important a word is in our database, so the words that appear less in our database have a higher score than words that appear more often. The free text search implements standard enhancements, such as stopword removal and term stemming for the English language. Since the annotations for the LSC dataset are generated automatically by content analysis tools, the natural variability in human text does not occur in the LSC collection. Hence, concepts such as *term frequency* weighting of terms are unlikely to play any significant role in the ranking process. However, the *inverse document frequency concept* from information retrieval is important in that it allows for the higher ranking of concepts and our use of TF-IDF ranking includes the inverse document frequency weighing in the IDF component.

This ranked list from the free-text search is filtered by the other data facets, such as time of day, day of week, or location, as is commonly employed for interactive retrieval systems [14]. The result is a ranked list of filtered images for presentation to the user.

One enhancement to the ranked list is that the context of the image is also considered in the scoring process. The previous three images and the following three images contribute (on a sliding scale) to the overall score of the main image. This is based on previous successful deployments in interactive video retrieval systems [1].

2.4 Result List Presentation

Upon submission of a query, the system generates a ranked list of images. The images are displayed as you can see in the *Figure 1* on the result-display panel to the right of the query panel. Rather than simply display a ranked list of images, the system aims to provide some context to the user by integrating two enhancements:

- Providing context to the main image by highlighting meta-data regarding the image, such as temporal, location and activity data.
- Highlighting the visual context of the image by borrowing from concepts previously used in interactive video retrieval systems [1], the visual context of each ranked image is displayed by the highlighting of 1 or 2 (1 shown in *Figure 1*) previously occurring and following images. Given the passive nature of lifelog data capture, a naive approach would be to simply select the previous occurring image(s) to be shown on screen. However, lifelog data tends to be very repetitive, so a judicious approach is used which only shows the previews and following images if they are sufficiently visually different from the ranked image in terms of the visual content and the concept annotations. We include this functionality because the development topics contained some temporal topics that require knowledge about the preceding or succeeding activities of the lifelogger.

Upon finding a potentially relevant image for a topic, the user selects an image which submits it to the LSC server for validation.

3 CONCLUSION

In this paper, we describe the data and the task from the LSC 2018 lifelog search challenge. We describe the prototype interactive search engine that we built for the challenge and we presented the interface and extensions to the basic search engine.

We have selected a number of facets for the query interface (as described above) and we are limited by the nature of the LSC collection, as well as the available screen estate for the query-entry process. Knowing that integrating more facets are likely to increase performance [10], future work will seek to enhance the types and richness of the facets available to the user. In addition, the implementation of free-text search in this first generation lifelog search engine does not include many of the standard text retrieval enhancements such as relevance feedback [12] or optimised ranking algorithms (e.g. BM25) [11]. For future work, consideration should be given to optimising the system components.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of Science Foundation Ireland (SFI) under grant number SFI/12/RC/2289.

REFERENCES

- [1] Eddie Cooke, Paul Ferguson, Georgina Gaughan, Cathal Gurrin, Gareth J.F. Jones, Herve Le Borgne, Hyowon Lee, Sean Marlow, Kieran McDonald, Mike McHugh, Noel Murphy, Noel E. O Connor, Neil O Hare, Sandra Rothwell, Peter Wilkins, and Alan F. Smeaton. 2004. TRECVID 2004 experiments in Dublin City University. *TRECVID 2004 experiments in Dublin City University* (2004). arXiv:1404.7828 <http://arxiv.org/abs/1404.7828>
- [2] Gabriel de Oliveira-Barra, X. Giró-i Nieto, Alejandro Cartas-Ayala, and Petia Radeva. 2016. LEMoRe: A Lifelog Engine for Moments Retrieval at the NTCIR-Lifelog LSAT Task. In *The 12th NTCIR Conference, Evaluation of Information Access Technologies*. National Institute of Informatics (NII), National Institute of Informatics (NII), Tokyo, Japan. [http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings12/NTCIR/toc_ntcir.html#\\$Lifelog](http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings12/NTCIR/toc_ntcir.html#$Lifelog)
- [3] Aiden R. Doherty, Kata Pauly-Takacs Niamh Caprani, Cathal Gurrin, Chris Moulin, Noel O Connor, and Alan F. Smeaton. 2012. Experiences of Aiding Autobiographical Memory Using the SenseCam. *Human-Computer Interaction* 27, 1-2 (2012), 151–174.
- [4] Aidan R. Doherty and Alan F. Smeaton. 2008. Automatically segmenting lifelog data into events. In *Ninth International IEEE Workshop on Image Analysis for Multimedia Interactive Services*. IEEE, 20–23.
- [5] Jim Gemmell, Gordon Bell, and Roger Lueder. 2006. MyLifeBits: a personal database for everything. *Commun. ACM* 49, 1 (2006), 88–95.
- [6] Cathal Gurrin, Hideo Joho, Frank Hopfgartner, Liting Zhou, and Rami Albalat. 2016. Overview of NTCIR-12 Lifelog Task. , 354–360 pages. <http://eprints.gla.ac.uk/131460/>
- [7] Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. 2014. LifeLogging: Personal Big Data. *Foundations and Trends in Information Retrieval* 8, 1 (2014), 1–125. <https://doi.org/10.1561/15000000033>
- [8] Bogdan Ionescu, Henning Müller, Mauricio Villegas, Helbert Arenas, Giulia Boato, Duc-Tien Dang-Nguyen, Yashin Dicente Cid, Carsten Eickhoff, Alba G. Seco de Herrera, Cathal Gurrin, Bayzidul Islam, Vassili Kovalev, Vitali Liauchuk, Josiane Mothe, Luca Piras, Michael Riegler, and Immanuel Schwall. 2017. Overview of ImageCLEF 2017: Information Extraction from Images. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, Gareth J.F. Jones, Séamus Lawless, Julio Gonzalo, Liadh Kelly, Lorraine Goeuriot, Thomas Mandl, Linda Cappellato, and Nicola Ferro (Eds.). Springer International Publishing, Cham, 315–337.
- [9] Jochen Meyer, Steven Simske, Katie Siek, Cathal Gurrin, and Hermie J Hermens. 2014. Beyond quantified self: data for wellbeing. In *Extended Abstracts on Human Factors in Computing Systems*. <http://dl.acm.org/citation.cfm?id=2560469>
- [10] Mor Naaman, Susumu Harada, QianYing Wang, Hector Garcia-Molina, and Andreas Paepcke. 2004. Context data in geo-referenced digital photo collections. In *ACM Multimedia*.
- [11] Stephen E. Robertson and Karen Sparck Jones. 1994. Simple, proven approaches to text retrieval. *University of Cambridge Technical Report* (1994). <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-356.pdf>
- [12] J. J. Rocchio. 1971. *Relevance Feedback in Information Retrieval*. Prentice Hall, Englewood, Cliffs, New Jersey. http://www.is.informatik.uni-duisburg.de/bib/docs/Rocchio_71.html
- [13] Jürgen Schmidhuber. 2014. Deep Learning in Neural Networks: An Overview. *CoRR* abs/1404.7828 (2014). arXiv:1404.7828 <http://arxiv.org/abs/1404.7828>
- [14] Daniel Tunkelang. 2009. Faceted Search. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 1, 1 (2009), 1–80. <https://doi.org/10.2200/S00190ED1V01Y200904ICR005> arXiv:https://doi.org/10.2200/S00190ED1V01Y200904ICR005